

Climate spectrum estimation in the presence of timescale errors

M. Mudelsee^{1,2}, D. Scholz^{3,4}, R. Röthlisberger¹, D. Fleitmann^{5,6}, A. Mangini³, and E. W. Wolff¹

¹British Antarctic Survey, Cambridge, UK

²Climate Risk Analysis, Hannover, Germany

³Heidelberg Academy of Sciences, Heidelberg, Germany

⁴Bristol Isotope Group, School of Geographical Sciences, University of Bristol, Bristol, UK

⁵Institute of Geological Sciences, University of Bern, Bern, Switzerland

⁶Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

Received: 3 December 2008 – Revised: 15 January 2009 – Accepted: 15 January 2009 – Published: 5 February 2009

Abstract. We introduce an algorithm (called REDFITmc2) for spectrum estimation in the presence of timescale errors. It is based on the Lomb-Scargle periodogram for unevenly spaced time series, in combination with the Welch's Overlapped Segment Averaging procedure, bootstrap bias correction and persistence estimation. The timescale errors are modelled parametrically and included in the simulations for determining (1) the upper levels of the spectrum of the red-noise AR(1) alternative and (2) the uncertainty of the frequency of a spectral peak. Application of REDFITmc2 to ice core and stalagmite records of palaeoclimate allowed a more realistic evaluation of spectral peaks than when ignoring this source of uncertainty. The results support qualitatively the intuition that stronger effects on the spectrum estimate (decreased detectability and increased frequency uncertainty) occur for higher frequencies. The surplus information brought by algorithm REDFITmc2 is that those effects are quantified. Regarding timescale construction, not only the fixpoints, dating errors and the functional form of the age-depth model play a role. Also the joint distribution of all time points (serial correlation, stratigraphic order) determines spectrum estimation.

1 Introduction

A classical decomposition of a climate process is into a trend and a noise component. Spectral analysis investigates the noise component. A Fourier transform into the frequency domain makes it possible to separate short-term from long-term

variations and to distinguish between cyclical forcing mechanisms of the climate system and broad-band resonances. Spectral analysis allows to learn about the climate physics.

The major task is to estimate the spectral density function. This poses more difficulties than, for example, linear regression (which may be used for trend estimation) because here we estimate a function and not just two regression parameters (intercept and slope). Spectral smoothing becomes therefore necessary, and this brings a tradeoff between estimation variance and frequency resolution.

In the case of evenly spaced records, the multitaper smoothing method achieves the optimal tradeoff (Thomson, 1982). The focus of this paper, however, is spectrum estimation for unevenly spaced records. Such a situation is ubiquitous in palaeoclimatology. In this case the method of choice is combining the Lomb-Scargle periodogram with Welch's Overlapped Segment Averaging (Welch, 1967; Lomb, 1976; Scargle, 1982; Schulz and Mudelsee, 2002), which yields an estimation in the time domain and avoids distortions caused by interpolation.

Bootstrap resampling enhances Lomb-Scargle methods by providing a bias correction. It supplies also a detection test for a spectral peak against realistic noise alternatives in form of an AR(1) model (which belongs to the wider class of “red noise” processes). We introduce a bootstrap algorithm, called REDFITmc2, which takes into account the effects of timescale uncertainties on detectability and frequency resolution. REDFITmc2 is a further development of the REDFIT algorithm (Schulz and Mudelsee, 2002), which assumes negligible timescale errors.

Palaeoclimate time series from two different archives serve to illustrate our concept. The first example is an ice core from Antarctica, which covers the past 800 ka and



Correspondence to: M. Mudelsee
(mudelsee@mudelsee.com)

constitutes the, at present, longest glacial archive. The second is a stalagmite from southern Arabia (Oman), which covers a large portion of the Holocene (last ~ 11 ka).

2 Method

Let $X(T)$ be a stationary climate process in continuous time, T . The one-sided non-normalized power spectral density function, $h(f)$, where $f \geq 0$ is frequency, can be defined (Priestley, 1981, Chapter 4 therein) via a truncation of $X(T)$ and a Fourier transformation. $h(f)$ integrates to $\text{VAR}[X(T)] = S^2$, the variance of $X(T)$. The power is a measure of the variability (“energy”) within a frequency interval $[f; f+df]$.

The spectral theory in the case of a process $X(i)$ in discrete time, $T(i)$, is similar (Priestley, 1981, Sect. 4.8.3 therein), except that the frequency range has an upper bound and the discrete Fourier transform is invoked to calculate $h(f)$.

The objective of spectrum analysis is to estimate $h(f)$ using a time series sample, $\{t(i), x(i)\}_{i=1}^n$. The time points, $t(i)$, may be unevenly spaced. n is the sample size. The spacing is given by $d(i) = t(i) - t(i-1)$, its average is denoted as \bar{d} . A spectral peak at a frequency, f' , corresponds to a period we denote as $T_{\text{period}} = 1/f'$.

2.1 Lomb-Scargle periodogram

An early spectrum estimator is Schuster’s (1898) periodogram, $I(f_j)$, which is evaluated at discrete frequency points, f_j . Scargle (1982) suggested for the case of uneven spacing a new version,

$$I_{\text{LS}}(f_j) = \bar{d} \cdot \left\{ \frac{\left[\sum_{i=1}^n X(i) \cos(2\pi f_j [T(i) - \tau_{\text{Lomb}}]) \right]^2}{\sum_{i=1}^n [\cos(2\pi f_j [T(i) - \tau_{\text{Lomb}}])]^2} + \frac{\left[\sum_{i=1}^n X(i) \sin(2\pi f_j [T(i) - \tau_{\text{Lomb}}]) \right]^2}{\sum_{i=1}^n [\sin(2\pi f_j [T(i) - \tau_{\text{Lomb}}])]^2} \right\}, \quad (1)$$

where Lomb’s (1976) time shift, τ_{Lomb} , is given via

$$\tan(4\pi f_j \tau_{\text{Lomb}}) = \frac{\sum_{i=1}^n \sin(4\pi f_j T(i))}{\sum_{i=1}^n \cos(4\pi f_j T(i))}. \quad (2)$$

For even spacing ($d(i) = d$), even n and frequency points $f_j = 1/(nd), \dots, 1/(2d)$, it follows that $\tau_{\text{Lomb}} = 0$ and $I_{\text{LS}}(f_j) = I(f_j)$. To calculate $I_{\text{LS}}(f_j)$ on the sample level, that is, using the time series sample, plug in $x(i)$ for $X(i)$ and $t(i)$ for $T(i)$.

Scargle’s objective behind introducing the Lomb-Scargle periodogram was that the probability distribution of $I_{\text{LS}}(f_j)$ should be equal to the distribution of $I(f_j)$. Scargle (1982, 1989) showed that this is so (chi-squared distribution) when $X(i)$ is a Gaussian white noise process (with mean zero and variance S^2), $X(i) = \mathcal{E}_{N(0, S^2)}(i)$.

2.2 Welch’s overlapped segment averaging

The periodogram is a suitable estimator of line spectra, where harmonic signal components appear as sharp peaks. To enhance the unfavourable inconsistency of the periodogram (i.e., the variance does not decrease with n increasing) when (mis-)applied to estimate continuous spectra, Welch (1967) advanced the idea of Bartlett (1950) to divide a time series $\{t(i), x(i)\}_{i=1}^n$ into different time segments, calculate the periodograms segment-wise and average them to obtain a reduced estimation variance. Welch (1967) allowed the segments to overlap (for example, by 50%), and the method is called “Welch’s Overlapped Segment Averaging” or WOSA procedure. Overlapping has the positive effect that the number of segments, and therefore the number of averaged periodograms, is increased.

The negative effect of using WOSA (number of segments, $n_{50} > 1$) is that the frequency points, where the periodograms are calculated, are spaced wider than for $n_{50} = 1$. More precisely, $\Delta f_j = (n_{50} + 1)/(2nd) > 1/(nd)$ for $n_{50} > 1$. This is the smoothing problem in the spectral domain, the tradeoff between spectral estimation variance and frequency resolution.

Instead of calculating the Lomb-Scargle periodogram at frequencies $f_j = 1/(n\bar{d}), 2/(n\bar{d}), \dots$, there is no hindrance to using a finer frequency grid. The oversampling factor (OFAC) (Schulz and Stattegger, 1997) gives the increase in frequency resolution. Oversampling is mainly for “cosmetic” reasons, to have a smoother curve and to be able to determine precisely the frequency of a spectral peak. Like interpolation in the time domain, oversampling does not generate new information and cannot solve the smoothing problem.

Welch (1967) showed also that tapering (weighting) the data points $X(i)$ within segments improves the spectrum estimate in terms of bias, variance and suppression of spurious peaks (sidelobes). The overview by Harris (1978) lists the properties of many taper functions, including the Welch taper (negative parabola) that is used in the present paper.

2.3 Bandwidth

The degrees of freedom of the chi-squared distribution of a Lomb-Scargle spectrum estimate based on WOSA with 50% overlap and Gaussian distributed $X(i)$ are

$$\nu = 2n_{50} / \left(1 + 2c^2 - 2c^2/n_{50} \right), \quad (3)$$

where $c \leq 0.5$ is a constant representing the taper. A uniform taper has $c = 0.5$, a Welch taper $c = 0.344$; further values are listed by Harris (1978).

The discrete Fourier transform of a pure harmonic process ($X(i)$ composed of sinus and cosinus terms with frequency f') has a peak at f' . The decay on the flanks of the peak to a value of $10^{-6/10} \approx 0.251$ times the maximum value defines a width in frequency, the 6-dB spectral bandwidth, B_s . This is a useful quantity for assessing the frequency resolution, how

good closely neighboured spectral peaks can theoretically be separated (Harris, 1978; Nuttall, 1981). The 6-dB bandwidth depends on n , n_{50} and the type of taper.

2.4 Bootstrap bias correction

If $X(i)$ is a red-noise process on an unevenly spaced timescale, perhaps with superimposed peaks in the frequency domain, the distribution of $I_{LS}(f_j)$ cannot be calculated analytically. Here, simulation methods can be used to infer the distributional properties of the Lomb-Scargle periodogram. Of particular interest is its bias (systematic error).

Monte Carlo experiments (Schulz and Mudelsee, 2002) reveal the bias of the Lomb-Scargle periodogram for an AR(1) process and uneven spacing. The “absolute bias”, of $I_{LS}(f_j)$ as an estimator of non-normalized power, $h(f)$, can be substantial. The bias disappears (i.e., becomes smaller than the “simulation noise”) for an AR(1) process and even spacing. Also in peak detection, where normalized power, $g(f)=h(f)/S^2$, is analysed, the Lomb-Scargle periodogram exhibits bias. That means, even if the normalization is known, the bias of the Lomb-Scargle periodogram is significant, and it is frequency-dependent.

Spectrum estimation for unevenly spaced time series can be enhanced by combining the WOSA procedure with the Lomb-Scargle periodogram (Schulz and Stattegger, 1997). A bias correction for such estimates was devised by Schulz and Mudelsee (2002). It uses a bootstrap simulation approach based on artificially generated AR(1) time series (“surrogate data”), for which the theoretical spectrum is known (Priestley, 1981). The frequency-dependent bias correction factor is calculated as the ratio of the theoretical AR(1) spectrum and the average Lomb-Scargle spectrum. The bias-corrected spectrum estimate is $\hat{h}'(f_j)$.

2.5 Peak detection by red-noise test

A hypothesis test helps to assess whether peaks (or lows) in the estimate of $h(f)$, denoted as $\hat{h}(f)$, are significant or not. Such information is for the climate time series analyst of major relevance because it helps to filter out the variability, to construct and test conceptual climate models – the absolute value of $h(f)$ is less important. The typical test performed in climate spectral analysis is of H_0 : “ $X(i)$ is an AR(1) process, with continuous, red spectrum”, the red-noise hypothesis, against H_1 : “ $X(i)$ has a mixed spectrum, with peak at $f_j=f'_j$ on a red-noise background.”

Schulz and Mudelsee (2002) devised the following red-noise test. The null distribution of $\hat{h}(f)$ is obtained by fitting an AR(1) process to $\{t(i), x(i)\}_{i=1}^n$, that is, estimating the persistence time, τ , followed by bootstrap resampling.

The process is given by

$$\begin{aligned} X(1) &= \mathcal{E}_{N(0, 1)}(1), \\ X(i) &= \exp\{-[T(i) - T(i-1)]/\tau\} \cdot X(i-1) \\ &\quad + \mathcal{E}_{N(0, 1-\exp\{-2[T(i)-T(i-1)]/\tau\})}(i), \\ i &= 2, \dots, n. \end{aligned} \quad (4)$$

The parameter τ defines the “equivalent autocorrelation coefficient”, a' , via $a' = \exp(-\bar{d}/\tau)$ for the case of uneven time spacing. For even spacing, a' equals a , the ordinary AR(1) autocorrelation parameter.

Even the estimation of a (even spacing) is not trivial because of estimation bias. For an AR(1) process with unknown mean, which is the usual case in climatology, the estimator of a is

$$\hat{a} = \frac{\sum_{i=2}^n [X(i) - \bar{X}] \cdot [X(i-1) - \bar{X}]}{\sum_{i=2}^n [X(i) - \bar{X}]^2}, \quad (5)$$

where $\bar{X} = \sum_{i=1}^n X(i)/n$ is the sample mean. The approximate expectation of \hat{a} is (Kendall, 1954)

$$E(\hat{a}) \simeq a - (1 + 3a) / (n - 1). \quad (6)$$

That means, \hat{a} underestimates a . Eq. (6) can be used to correct for the negative bias. However, Eq. (6) is valid only for a not too large (Kendall, 1954). A bias formula applicable also to large values (above, say, 0.9) of a has yet only been derived for the simpler case of known mean (Mudelsee, 2001).

Regarding the estimation of τ (uneven spacing), Monte Carlo simulations (Mudelsee, M.: Climate Time Series Analysis: Classical Statistical and Bootstrap Methods, Springer, Heidelberg, manuscript in preparation) show that the least-squares estimator of $a' = \exp(-\bar{d}/\tau)$ has a bias similar in size to the bias of \hat{a} . Mudelsee (2002) presented a least-squares estimation procedure of τ with bias correction and bootstrap percentile confidence intervals. Our conclusions for the practice of persistence time estimation (Mudelsee, 2002) are as follows.

- Large data sizes and not too large values of τ make bias correction via the application of Eq. (6) to a' accurate,
- It may in practice be difficult to show statistically significant changes of τ over short time intervals and therefore be often advisable to assume a constant persistence time. We use this approach because own analyses (not shown) with the ice core records (Sect. 3.1) divided into sub-intervals, do not indicate significant changes in τ . An option that may apply to a situation where the physics of the sampled climate system indicates a change in the persistence properties and where also very long time series are available, may be to consider a model that divides the time interval into sub-intervals with specific persistence times. Divine et al. (2008)

have recently suggested a similar model (for even spacing), accompanied by a maximum likelihood estimation procedure.

Resampling is done (Schulz and Mudelsee, 2002) via the surrogate data approach. Instead of (or in addition to) performing bootstrap resampling for deriving the null distribution, one may calculate upper (lower) bounds via the chi-squared distribution (Schulz and Mudelsee, 2002). Of practical relevance in climatology, although conventionally ignored, should be not only peaks but also lows in $\hat{h}(f)$, frequency intervals where less power than under an AR(1) hypothesis resides. Also the information about the absence of power in a frequency range is helpful; it may, for example, support an argumentation that a certain mechanism does *not* play a role for generating an observed climate phenomenon.

So far we have assumed as a realistic noise alternative an AR(1) process with positive persistence time τ (uneven spacing) or positive autocorrelation parameter a (even spacing). This corresponds to a “red” spectrum, with more power for lower than for higher frequencies. Fisher et al. (1985) and Andersen et al. (2006) describe the interesting observation that postdepositional effects (wind erosion) at a site of an ice core may lead to “blue” noise, with more power for higher frequencies. This would make, however, the testing of the spectrum alternative problematic. Uneven spacing does not allow to embed a discrete-time, blue-noise process within a continuous-time, blue-noise process (Robinson, 1977; Chan and Tong, 1987). This means that for uneven spacing it is not possible to infer uniquely the parameters of the underlying continuous-time process. Embedding is an important property because it allows a foundation on physics, which works in continuous time (differential equations). A practical approach to ice core spectrum estimation may be to use extra caution in the interpretation of high-frequency peaks.

2.6 Multiple tests

Assume for convenience even data size, even spacing and no oversampling. If the red-noise hypothesis test for the existence of a spectral peak is to be carried out for one single, pre-defined frequency $f'_j \in \{f'_j\}_{j=1}^{n/2}$, then selection of the $100(1-\alpha)$ th percentage point of the red-noise distribution leads to a one-sided hypothesis test with a P-value equal to α . If, what is usually the case, the test is multiple, that means, it is to be carried out for many (if not all) frequencies from the set $\{f'_j\}_{j=1}^{n/2}$, then a higher frequency-point-wise confidence level, $(1-\alpha')$ with $\alpha' < \alpha$, has to be employed to yield an overall P-value of α . If a test is performed multiple times, it becomes more likely to find a significant single result.

One may define a “maximum effective number of test frequencies”, M , via the overall prescribed P-value: $(1-\alpha')^M = 1-\alpha$. For small α and large M this leads to $\alpha' \approx \alpha/M$. The effective number of frequencies refers to a hypothetical situation where M frequencies f'_j are tested and

the spectrum estimates $\hat{h}(f'_j)$, a set of size M , are mutually independent. For the simple case of even data size, even spacing, Gaussian distributional shape and periodogram estimation, independence is fulfilled and the maximum number is $M=n/2$. If n is odd (other setting unchanged), $M=(n-1)/2$. Also if the Gaussian assumption is violated not too strongly, the effects on M should be negligible.

Stronger influence on M can have uneven spacing with Lomb-Scargle periodogram estimation (no WOSA). Because the periodogram estimates are then no longer independent, M is reduced. Horne and Baliunas (1986) and VanDongen et al. (1997) studied the effects by means of Monte Carlo simulations. If the $\{t(i)\}_{i=1}^n$ are more or less uniformly distributed, the approximation $M \approx n/2$ is still acceptable. This formula may also be applied to series where the timescale is even with the exception of a few missing observations. However, if the time points are highly clustered in time, one should not use the number of points, n , but rather the number of clusters, to determine M (VanDongen et al., 1997). The effects of segmenting (WOSA) on M with Lomb-Scargle or ordinary periodogram estimation (no tapering) can be taken into account by using instead of n the number of points per segment: $M = \text{NINT}[n/(n_{50}+1)]$, see Schulz and Mudelsee (2002). NINT is the nearest integer function. The effects of tapering could in principle be studied by means of Monte Carlo simulations. Restricting the frequency range where to study peaks is an other way to reduce M .

What practical conclusions can be drawn for peak detection in climate spectra? A typical situation is an unevenly spaced timescale without strong clustering, and where the researcher is interested also in the longer periods of variations recorded by the time series. Here, Lomb-Scargle periodogram estimation with tapering, WOSA and n_{50} not too high (less than, say, 10) is an option. To have more reliability in the low-frequency spectrum portion, one decides to follow a rule of thumb (Bendat and Piersol, 1986) and requires at least two cycles per segment, that is, one sets the minimum test frequency f_j equal to $[(2n_{50})/(n\bar{d})]$. This also reduces M . Regarding the high-frequency spectrum portion, theoretically the uneven spacing allows inferences also for frequencies above $1/(2\bar{d})$, see Scargle (1982). On the other hand, an archive may a priori be known not to preserve a high-frequency signal, for example a marine sediment core affected by bioturbation. Then it would make sense to ignore a part of the high frequencies, leading to a further reduction of M .

2.7 Timescale errors

In palaeoclimatology the timescale is usually expected to have uncertainties. The time assigned to a sample, $T(i)$, determined by dating and constructing an age-depth curve, is expected to deviate from the true time value, $T_{\text{true}}(i)$. In such cases we write the measured times as

$$T(i) = T_{\text{true}}(i) + T_{\text{noise}}(i), \quad (7)$$

$i=1, \dots, n$, where $T_{\text{noise}}(i)$ is the error owing to age uncertainties. Equivalently, the spacing, $d(i)$, has an error component.

To recapitulate, $T(i)$ is a sequence of random variables representing the observed time of a sequence of samples, that is, $T(i)$ is a discrete-time stochastic process (“process level”). The sequence $t(i)$ are the actually observed time values (“sample level”), a realization of the $T(i)$. The function $T_{\text{true}}(i)$ is the true time – uniquely defined, but unknown numbers. Because of dating errors we cannot expect that $T(i)$ equals $T_{\text{true}}(i)$. Hence, there is a noise component, $T_{\text{noise}}(i)$. The noise process is not required to have zero mean. This allows for measurement bias, that is, systematic over- or underestimations. Eq. (7) is a rather general description of uncertain timescales; it may be used for various combinations of climate archives and dating techniques.

Timescale errors lead to a distortion of the estimated spectrum. Two effects are expected:

1. reductions of significance (detectability) of peaks compared to a situation with exact timescale and
2. increases of frequency uncertainty for a detected spectral peak.

Moore and Thomson (1991) and Thomson and Robinson (1996) studied the influence of a “jittered” spacing on the process level. The simple case of independent Gaussian jitter,

$$d(i) = d + \mathcal{E}_{N(0, \delta_d^2)}(i), \quad (8)$$

is analytically tractable. δ_d is the standard deviation of the spacing error distribution. Its effect on the true continuous-time spectrum, $h(f)$, amounts (Moore and Thomson, 1991; Wunsch, 2000) to a multiplication by a frequency-dependent factor:

$$h_{\text{distort}}(f) = \exp\left(-4\pi^2 \delta_d^2 f^2\right) \cdot h(f) + c_0, \quad (9)$$

where the constant c_0 serves to give the distorted spectrum, $h_{\text{distort}}(f)$, the nominal area of S^2 . This means, timescale errors in the form of independent jitter add white noise (c_0). As a result, spectral peaks have a reduced detectability.

Several assumptions went into the derivation of Eq. (9) by Moore and Thomson (1991), this limits its applicability to the practice of climate spectrum estimation.

- No aliasing ($h(f)=0$ for $f > f_{\text{Ny}}=1/(2d)$). This may in practice be violated to some degree, and lead to power “folded” back into the interval $[0; f_{\text{Ny}}]$ and, possibly, to spurious spectral peaks. In addition, for unevenly spaced time series the Nyquist frequency is not well defined.

- Independent jitter. This is not realistic for many records (e.g., from ice or sediment cores). Moore and Thomson (1991) study AR(1) dependence in the jitter equation (Eq. 8), finding potential for larger effects on the spectrum if the dependence is strong. Still it is questionable how relevant the AR(1) jitter model is. Ice core data could exhibit heteroscedastic jitter owing to compaction. Timescales derived from layer counting might be better described by means of a random walk (i.e., a nonstationary AR(1) process with parameter $a=1$) than by a jitter model. The argument is as follows. If a layer is miscounted (e.g., missed), then this error propagates without “damping” to the next layer to be counted. It is rather difficult to obtain analytical results in such cases.
- Gaussian jitter distribution. This assumption is not fulfilled without imposing a constraint to guarantee a monotonic age-depth curve. (Moore and Thomson, 1991 had the purpose to study other data, in the spatial domain, where no such constraint is required.)
- Process level. The mentioned paper does not study the spectrum estimators on the sample level, in particular, Lomb-Scargle estimation.

Based on the limited relevance of available analytical results on the effects of realistic types of timescale errors on spectrum estimates (Lomb-Scargle), we suggest a novel numerical algorithm. It is based on simulations of the timescale. One mode of it quantifies the reduced detectability, the other the increased frequency uncertainty.

2.7.1 Algorithm REDFITmc2, mode “a”: peak detection

Algorithm REDFITmc2 adapts the REDFIT algorithm (Schulz and Mudelsee, 2002) to take into account timescale errors. Mode “a” of the algorithm performs peak detection.

1. Time series, $\{t(i), x(i)\}_{i=1}^n$.
2. Bias-corrected Lomb-Scargle spectrum estimate (Schulz and Mudelsee, 2002), $\hat{h}'(f_j)$.
3. Estimated, bias-corrected persistence time (Mudelsee, 2002), $\hat{\tau}'$.
4. Determine area, $A_{\hat{h}'}$, under spectrum within $[0; 1/(2\hat{d})]$.
5. Generate AR(1) data, $\{t(i), x_{\text{AR}(1)}^*(i)\}_{i=1}^n$ after Eq. (4).
6. Use timescale model (an implementation of Eq. 7) to simulate times, $\{t^*(i)\}_{i=1}^n$.
7. Bias-corrected Lomb-Scargle spectrum estimate, $\hat{h}'^{*b}(f_j)$, calculated from $\{t^*(i), x_{\text{AR}(1)}^*(i)\}_{i=1}^n$. Spectrum estimate subsequently scaled to area $A_{\hat{h}'}$. b counts how often Steps 5 to 7 are called.

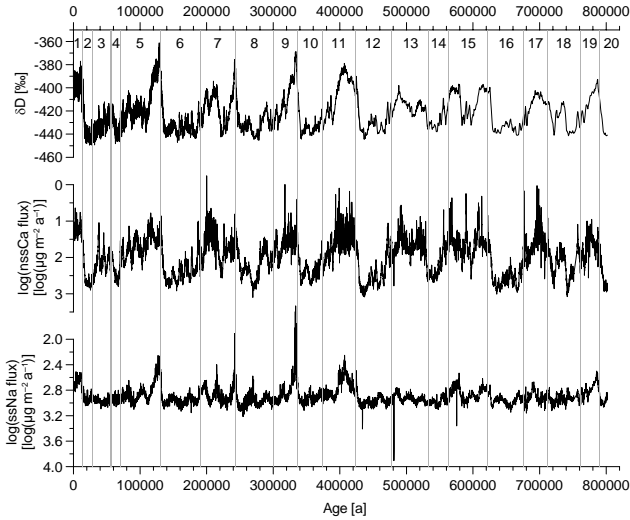


Fig. 1. Ice core EDC, data; δD , 55-cm resolution data ($n=5785$, $\bar{d}=139$ a); $\log(\text{nssCa flux})$, 100-a means ($n=7880$, $\bar{d}=102$ a) and $\log(\text{ssNa flux})$, 100-a means ($n=7891$, $\bar{d}=102$ a). A small number of the 100-a intervals have no data. Timescale is EDC3 (Parrenin et al., 2007). Vertical grey lines and numbers indicate marine isotope stages (Lisiecki and Raymo, 2005).

8. Go to Step 5 until $b=B$ replications exist.
9. Test at each f_j whether $\hat{h}'(f_j)$ exceeds a pre-defined upper percentile of $\{\hat{h}'^{*b}(f_j)\}_{b=1}^B$.

The selection of B depends on the size of the percentile; higher percentiles require higher values of B .

The sets of frequencies f_j at Steps 2 and 7 are identical.

Taking into account that the test is usually performed not only at one single frequency but at $f_j=0, 1/(n\bar{d}), \dots, 1/(2\bar{d})$ requires to adjust the significance level for a single test. See Sect. 2.6.

2.7.2 Algorithm REDFITmc2, mode “b”: frequency uncertainty

Mode “b” of the algorithm quantifies the effect of timescale errors on the uncertainty of the frequency of a spectral peak.

1. Time series, $\{t(i), x(i)\}_{i=1}^n$.
2. Bias-corrected Lomb-Scargle spectrum estimate (Schulz and Mudelsee, 2002), $\hat{h}'(f_j)$.
3. Estimated, bias-corrected persistence time (Mudelsee, 2002), $\hat{\tau}'$.
4. Determine area, $A_{\hat{h}'}$, under spectrum within $[0; 1/(2\bar{d})]$.

5. Spectral peak at frequency f'_j , area under peak,

$$\int_{f'_j-0.5B_s}^{f'_j+0.5B_s} \hat{h}'(f) df = \gamma \cdot A_{\hat{h}'}$$

6. Generate AR(1) data, $\{t(i), x_{\text{AR}(1)}^*(i)\}_{i=1}^n$ after Eq. (4).
7. Generate sinusoidal data, $\{t(i), x_{\text{sin}}^*(i)\}_{i=1}^n$, with $x_{\text{sin}}^*(i) = (2\gamma)^{1/2} \sin(2\pi f'_j t(i))$.
8. Mix series: $x^*(i) = (1-\gamma)^{1/2} x_{\text{AR}(1)}^*(i) + x_{\text{sin}}^*(i)$. Note that $\text{VAR}[X^*(i)] = (1-\gamma) + \text{VAR}[X_{\text{sin}}^*(i)] \approx 1$.
9. Use timescale model to resample times, $\{t^*(i)\}_{i=1}^n$.
10. Bias-corrected Lomb-Scargle spectrum estimate, $\hat{h}'^{*b}(f_j)$, calculated from $\{t^*(i), x^*(i)\}_{i=1}^n$, scaled to area $A_{\hat{h}'}$ under peak at f_j^* . b , counter.
11. Go to Step 6 until $b=B$ (usually $B=2000$ is sufficient (Efron and Tibshirani, 1993)) versions of f_j^* exist.
12. Calculate standard error, $\text{se}_{f'_j}$, by taking the standard deviation of $\{f_j^{*b}\}_{b=1}^B$.

3 Examples

We illustrate spectrum estimation using two types of palaeoclimate archives: ice cores and speleothems. We describe shortly the records (proxy variables) and their timescales. The results of the application of the REDFITmc2 algorithm are discussed with a focus on statistical aspects. Climatic interpretations shall be presented elsewhere.

3.1 Ice core: EPICA Dome C (Antarctica)

The European Project for Ice Coring in Antarctica (EPICA) core from the Dome C site (75° S, 123° E) has a length of ~ 3260 m and covers the past $\sim 800\,000$ a (EPICA Community Members, 2006; Wolff et al., 2006; Jouzel et al., 2007). We refer to this as the EDC core.

3.1.1 Data

The deuterium isotope (δD) time series (Jouzel et al., 2007) records variations in air temperature at the inversion height over the EDC site. This proxy may indicate, at a reduced accuracy, also the temperature changes over the entire southern hemisphere. The non-sea-salt calcium (nssCa) flux record (Wolff et al., 2006; Röthlisberger et al., 2008) documents variations in the climate variable dust (transported from South America to Antarctica). The sea-salt sodium (ssNa) flux (Wolff et al., 2006; Röthlisberger et al., 2008) series is a proxy for changes in the extent of sea ice around Antarctica. The nssCa flux and ssNa flux records have rather strongly

skewed data distributions. We employ the logarithmic transformation (Röthlisberger et al., 2008) to bring the distributions closer to a symmetric/normal shape.

All three climate variables display long-term variations during the late Pleistocene (Fig. 1) and belong potentially to the set of variables governing the ice-age climate. Spectrum estimates may shed further light to the current understanding (Raymo and Huybers, 2008; Röthlisberger et al., 2008) of glacial–interglacial transitions. We follow Jouzel et al. (2007) and divide the interval into an earlier (400–800 ka) and a later (0–400 ka) interval.

3.1.2 Spectra, without timescale errors

First, we present the raw periodograms (Fig. 2), then the climate spectra calculated with timescale errors ignored (Fig. 3) and compare finally (Sect. 3.1.4) some of these results with those resulting from taking timescale errors into account.

In the δD spectrum for the earlier interval (400–800 ka) (Fig. 3b), from the Milankovitch periods only one peak appears to exist, at the period of obliquity variations (41 ka). Jouzel et al. (2007) find this and also an other peak, at a period of roughly 100 ka. For the later interval (0–400 ka) there are three peaks in the Lomb–Scargle spectrum (Fig. 3a), at the periods 109 ka, 40 ka and 23 ka. This agrees with the findings of Jouzel et al. (2007).

The question of the statistical significance of the Milankovitch peaks of the δD spectra is deferred to Sect. 3.1.4. Here we remark that the peaks at 64 ka and 30 ka period in the unsmoothed periodogram for δD (Fig. 2a), do not appear in the (more reliable) smoothed spectrum estimate (Fig. 3a, b). The reason could be that the shortness of the 400-ka intervals leads to a large bandwidth, B_s , which in turn prevents detection and separability of those peaks. An other reason is that their appearance in the unsmoothed periodogram is spurious owing to the large estimation variance.

The $\log(\text{nssCa flux})$ record displays remarkably “clean” spectra (Fig. 3c, d), with power concentrated in two frequency bands: approximately 100–109 ka and approximately 41–50 ka. The $\log(\text{ssNa flux})$ spectra (Fig. 3e, f), on the other hand, show more peaks that lie above the upper red-noise levels.

We also show spectra for the Holocene (Fig. 4). These were calculated using records of very high temporal resolution (δD , 55-cm resolution; nssCa flux and ssNa flux, 2-a downsampled). A number of peaks appear, mainly in the centennial band.

3.1.3 Timescale

The “official” timescale, called EDC3, was produced (Parrenin et al., 2007) by employing a snow accumulation and ice-flow model and invoking constraints in form of a set of independent age markers. The markers are given, on the one hand, by events such as a reversal of the Earth’s magnetic

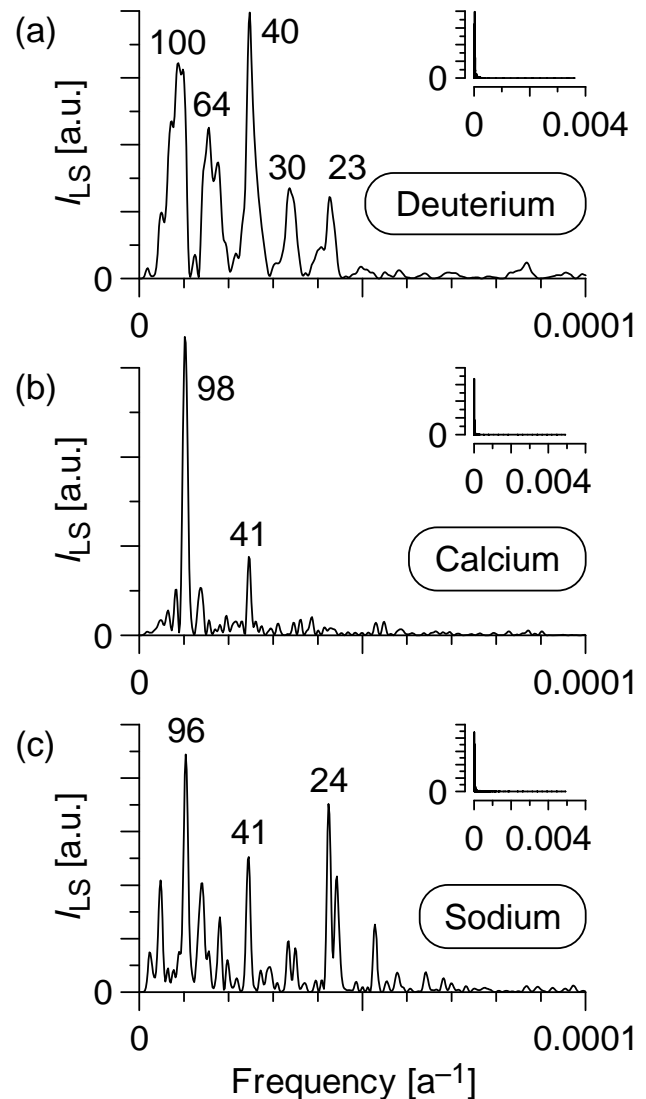


Fig. 2. Periodograms, EDC records: (a) δD , (b) $\log(\text{nssCa flux})$ and (c) $\log(\text{ssNa flux})$. Peaks are labelled with period values (ka). Insets show full frequency intervals. Periodograms were calculated on the records shown in Fig. 1 subsequent to linear detrending. Lomb–Scargle parameters: OFAC=6, $n_{50}=1$ and uniform taper. a.u., arbitrary units.

field or a volcanic eruption that has been recorded by an absolutely dated palaeoclimatic archive. On the other hand, pattern matching of the proxy climate series with insolation series produced an other kind of absolutely dated fixpoints.

Like any palaeoclimatic timescale, the EDC3 timescale relies on the validness of the assumptions made (e.g., model formulation, model parameter values) and the precision of numerical input values (e.g., absolute dates of markers). The EDC3 timescale is, thus, an estimate of the true, but unknown timescale. Parrenin et al. (2007, Table 1 therein) give as

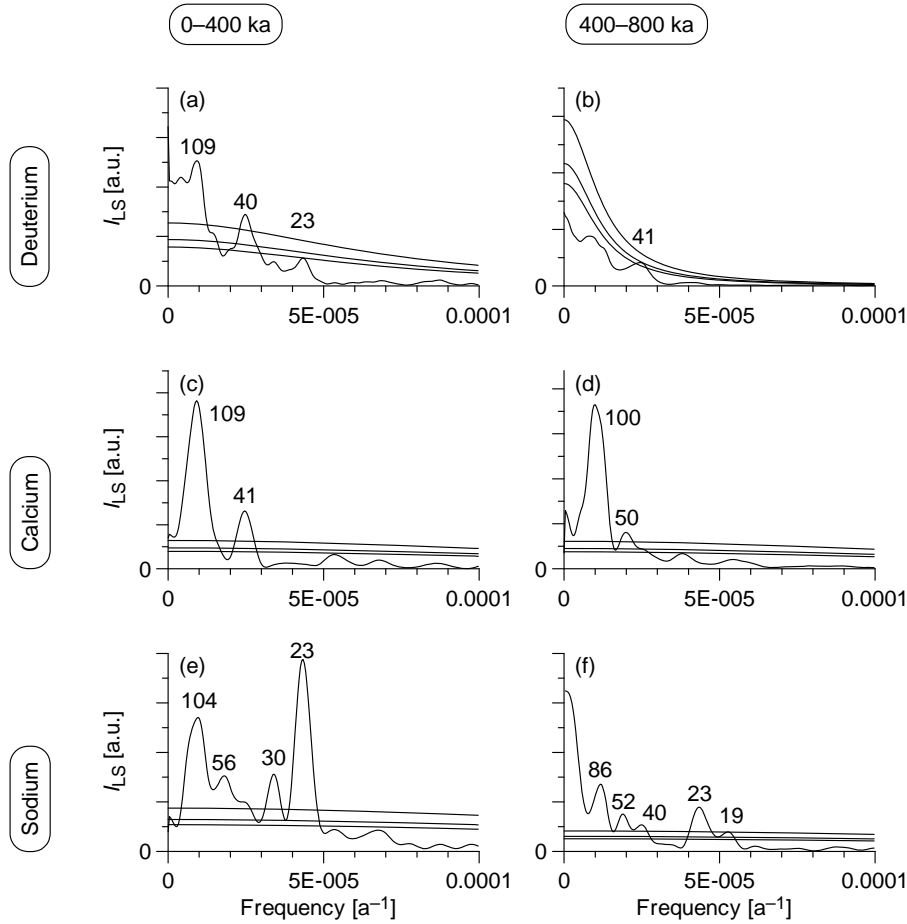


Fig. 3. Spectra, EDC records; (a) δD , interval 0–400 ka; (b) δD , interval 400–800 ka; (c) $\log(\text{nssCa flux})$, interval 0–400 ka; (d) $\log(\text{nssCa flux})$, interval 400–800 ka; (e) $\log(\text{ssNa flux})$, interval 0–400 ka; and (f) $\log(\text{ssNa flux})$, interval 400–800 ka. Spectra were calculated using the Lomb-Scargle method with WOSA (OFAC=12, $n_{50}=3$, Welch taper, $B_s \approx 8 \cdot 10^{-6} \text{a}^{-1}$) and bias correction ($B=2000$). Peaks are labelled with period values (ka). Each panel shows upper red-noise levels (smooth lines, from top: 99%, 95% and 90%) from a chi-squared distribution for an AR(1) process with following equivalent autocorrelation coefficients: (a, b) $a'=0.965$, (c, d) $a'=0.904$ and (e, f) $a'=0.868$. a.u., arbitrary units.

error estimate ($1-\sigma$ standard deviation) values of 4000 a for $T \lesssim 400\,000$ a and 6000 a for $T \gtrsim 400\,000$ a.

The application of the REDFITmc2 algorithm requires a strictly monotonically increasing timescale. We utilized the regression model of Heegaard et al. (2005) in a Bayesian manner to produce copies of $\{t^*(i)\}_{i=1}^n$ by means of the posterior age distributions. The input to this method are the EDC3 dating fixpoints and their errors (Parrenin et al., 2007, Table 1 therein). The same method has been recently used for timescale construction on speleothem data (Spötl et al., 2008).

A summary of the Bayesian method is as follows. Owing to their finite precision, the determined fixpoint ages may overlap within the dating error bars. This may result in age reversals in the simulated fixpoint timescale. To avoid such reversals, we used a version of the Gibbs sampler, which is a

Markov-chain Monte Carlo implementation (Buck and Millard, 2004). In a first step, an initial fixpoint timescale is simulated by drawing random numbers from the age distribution of each individual fixpoint. Then a sample age for the youngest fixpoint (Parrenin et al., 2007, Table 1 therein) is obtained taking into account (1) its age distribution and (2) the condition that it is younger than all other fixpoints. If that condition is violated, the age-simulation is repeated until the stratigraphic constraint is fulfilled. Similarly, the second fixpoint age of the section is simulated using its age distribution and the condition that it is older than the fixpoint above but younger than all other samples. This simulation is performed for all fixpoint ages of the section, which completes one iteration of the Gibbs sampler and results in a strictly monotonically increasing sequence of simulated fixpoints. In the next step we applied the Heegaard et al. (2005) nonparametric

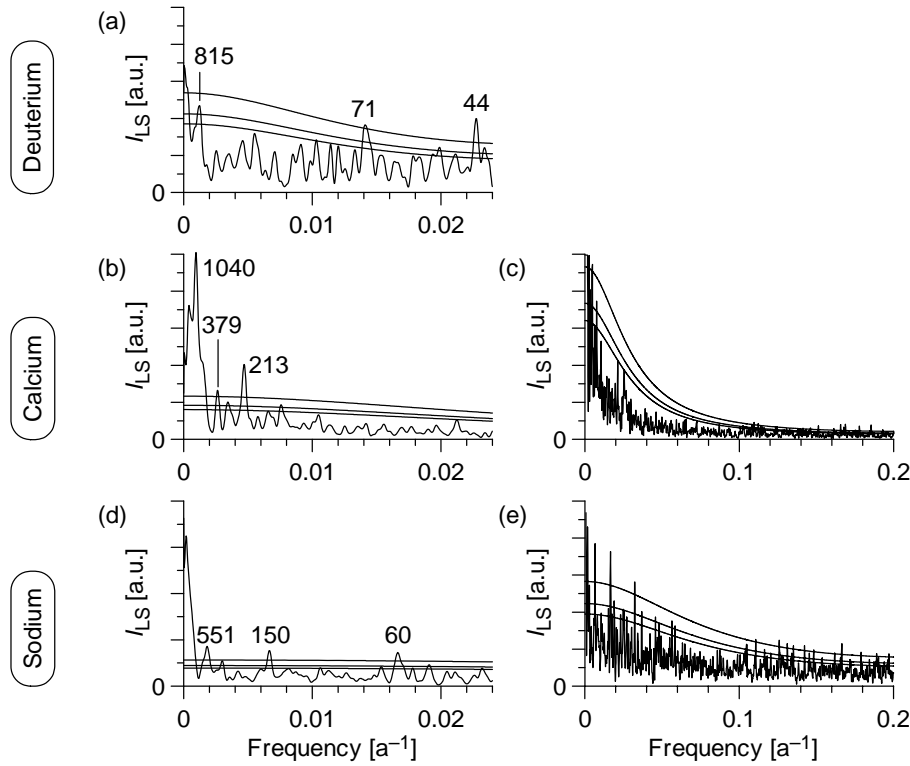


Fig. 4. Spectra, EDC records, Holocene interval (0–11 ka); (a) δD ; (b, c) $\log(\text{nssCa flux})$ and (d, e) $\log(\text{ssNa flux})$. Spectra were calculated on the 55-cm (a) or 2-a-downsampled (b–e) records. Parameters as in Fig. 3, with following changes: $n_{50}=6$, $B_s \approx 5 \cdot 10^{-4} \text{a}^{-1}$ and (a) $a'=0.181$, (b, c) $a'=0.648$ and (d, e) $a'=0.311$. Peaks are labelled with period values (years). a.u., arbitrary units.

regression model to the simulated fixpoints, providing a simulated timescale for the ice core proxy data. It may occur that the timescale produced by the Heegaard et al. (2005) algorithm does not strictly monotonically increase despite the simulated fixpoints do increase. (Such cases were rare in the case of the EDC simulations.) Thus, in a last step every simulated timescale, $\{t^*(i)\}_{i=1}^n$, is tested for apparent age reversals, and timescales with age reversals are discarded.

We apply the Bayesian method here to the ice core records for heuristic reasons. Other approaches for constructing $\{t^*(i)\}_{i=1}^n$ can be conceived, for example, piecewise-linear least squares regression (Sect. 3.2.2). The similarity between that (which may be labelled “frequentist”) and the “Bayesian” approach described in the preceding paragraph emphasizes the notion that the crucial point is not from which statistical school the method comes but that it reflects the underlying physics of the sampling of the climate archive. A further option is a “perturbed” version of physical timescale modelling (Parrenin et al., 2007). In the latter method, the B model runs are performed with model parameters not fixed but randomly drawn for each run.

3.1.4 Spectra, with timescale errors

The δD spectrum for the interval from 0 to 400 ka shows that the 40 ka obliquity cycle is significant also when taking into account timescale errors (Fig. 5a). The overall confidence level depends on the number of test frequencies, M . If we adopt a position that investigates only manifestations of three Milankovitch cycles (eccentricity, obliquity and precession) in the EDC records, then $M=3$ and the overall confidence level of the peak is approximately equal to $1 - (1 - 99\%) \times M = 97\%$ or higher (Fig. 5a). The uncertainty of the estimated peak frequency, se_{f_j} , is of the same order as the half bandwidth.

Also the significance of the EDC δD cycle at $T_{\text{period}}=109 \text{ ka}$ is robust against errors in the timescale (Fig. 5a). The major uncertainty associated with this cycle is the large bandwidth, which in turn results from the relative shortness of the selected time interval.

The significance of the precession cycle ($T_{\text{period}}=23 \text{ ka}$) in δD is already not high when ignoring timescale errors (Fig. 3a). We cannot therefore expect elevated significance levels in the more realistic test (obtained using t^*) and do not show that frequency range in the plot (Fig. 5a). It would be premature to conclude that a precession component does not

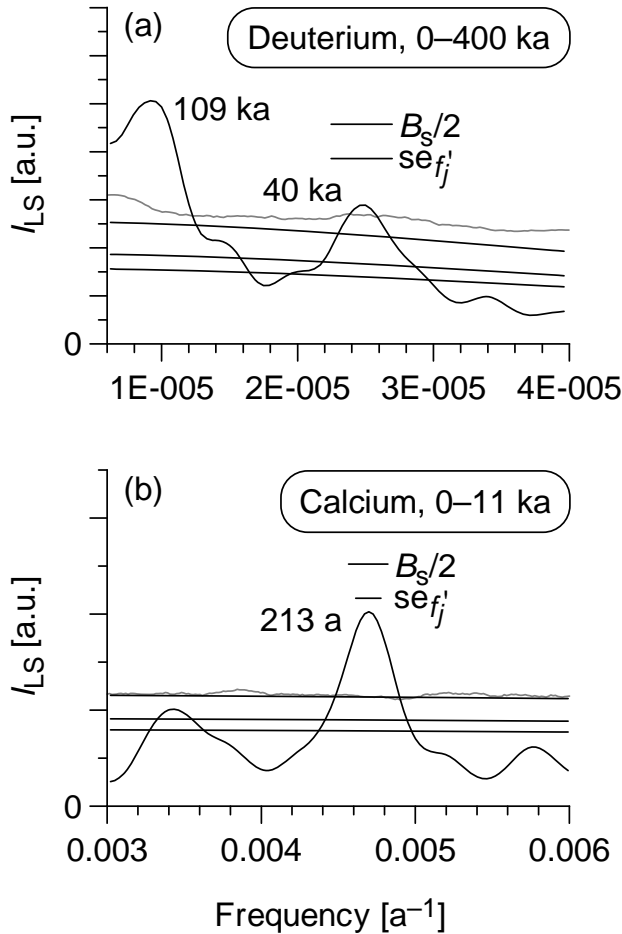


Fig. 5. Spectra, EDC records, effects of timescale uncertainties on spectral peaks; (a) δD , interval 0–400 ka, $T_{\text{period}}=109$ ka and 40 ka; (b) $\log(\text{nssCa flux})$, interval 0–11 ka, $T_{\text{period}}=213$ a. Spectrum estimation parameters are identical to those in Figs. 3a and 4b, respectively. Smooth solid black lines show (from bottom) upper 90%, 95% and 99% chi-squared bounds for AR(1) red-noise hypotheses. The influence of timescale uncertainties of the EDC record (Sect. 3.1.3) is studied using algorithm REDFITmc2. The 99% bootstrap bounds (grey, wiggly lines in a and b) are obtained from mode “a” (Sect. 2.7.1) with $B=10000$. The frequency uncertainties $se_{f_j'}$ (horizontal bars) due to timescale errors (mode “b”, Sect. 2.7.1), calculated with $B=10000$, are compared with the half of the bandwidth. (In (a), the value of $se_{f_j'}$ is for $T_{\text{period}}=109$ ka expected to be clearly smaller than $B_S/2$, and no calculations are performed.)

exist in Late Pleistocene temperature variations over Antarctica – making such a rather general statement would require the analysis of the large body of previous evidence (multiple records and ice cores) analysed using multiple statistical tests (Sect. 2.6). Such tests had to evaluate also the degree of circular reasoning invoked by astronomically tuning the palaeoclimate timescales. The less general statement that obliquity

has a stronger power than precession (temperature variations, 0–400 ka), on the other hand, appears to be supported well by the spectral analyses shown here.

The EDC $\log(\text{nssCa flux})$ record in the Holocene has a clear signal at $T_{\text{period}}=213$ a (Fig. 4b) that is significant at a high level also when taking into account timescale errors (Fig. 5b). These errors do not largely increase the accuracy of T_{period} . The reason for the small effects is that the EDC timescale is rather accurately known for the Holocene interval. Parrenin et al. (2007, Table 1 therein) give timescale errors ranging from 1–6 a in the latest part (last 800 a) to 50–180 a in the earlier part of the Holocene. On basis of the height of the peak, we suppose that the significance of the 213-year cycle is valid also when taking into account the multiplicity of the test (no calculations performed with higher percentiles).

There exists a considerable body of proxy evidence (Stuiver and Braziunas, 1993) for centennial solar activity variations during the Holocene; the cycle with a period of somewhat above 200 years is named after Hessel de Vries and Hans Suess. The coincidence of this solar cycle with the peak in the spectrum of $\log(\text{nssCa flux})$ should not be interpreted as a proof of solar forcing of dust flux variations in the southern hemisphere during the Holocene. Rather, this finding may encourage to explore such relationships using physical climate models.

3.2 Stalagmite: Q5 (Oman)

Stalagmite Q5 from the Qunf cave (17° N, 123° E) in Oman has a length of nearly 1 m and covers the interval from 10 300 a to 2740 a before the present (B.P.), see Fleitmann et al. (2003). Note that “present” is set to A.D. 1950.

3.2.1 Data

The oxygen isotope ($\delta^{18}\text{O}$) time series (Fleitmann et al., 2003) is a proxy mainly for changes in the intensity of Indian Ocean monsoonal rainfall (Neff et al., 2001; Fleitmann et al., 2003, 2007): high $\delta^{18}\text{O}$ values correspond to low rainfall amounts.

The Q5 $\delta^{18}\text{O}$ record (Fig. 6) displays certain long-term (trend) features: the onset of the Indian Ocean monsoon in the Early Holocene, a high (“optimum”) level until ~ 7200 a B.P. and a subsequent gradual decrease in rainfall intensity, which is thought to be related to the change in solar insolation (Fleitmann et al., 2003). The trend is punctuated by events of extreme dryness (heavy $\delta^{18}\text{O}$); the most prominent are the 8.2-ka and the 9.2-ka events (Fleitmann et al., 2008). We focus here on the interval [2740 a; 8000 a], that means, the time span after those events. The fitted trend was subtracted prior to spectrum estimation to extract the trend-stationary component. (REFITmc2 subtracts, as its predecessor REDFIT (Schulz and Mudelsee, 2002), segment-wise a linear fit).

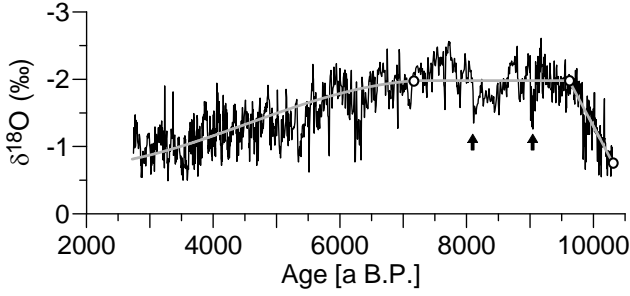


Fig. 6. Stalagmite Q5, data and trend. The trend (grey line) is composed of a function with change-points (open circles) in the early part, fitted by least-squares and a brute-force search (Mudelsee, 2000), and a sinusoid in the late part, fitted by ordinary least-squares. Spectral analysis focuses on the period after the 8.2-ka and 9.2-ka cold/dry extremes (arrows) (Fleitmann et al., 2008). This time interval [2740 a; 8000 a] has $n=973$ and $\bar{d}=5.4$ a.

3.2.2 Timescale

Fleitmann et al. (2003, Table S1 therein) contains the U/Th dates, $t_{\text{date}}(j)$, with standard errors, $s_{\text{date}}(j)$, made at stalagmite depth $z_{\text{date}}(j)$; $j=1, \dots, n_{\text{date}}$. The number of dating points required for bracketing the [2740 a; 8000 a] interval, is $n_{\text{date}}=11$, the average of $s_{\text{date}}(j)$ is ~ 70 a. Fleitmann et al. (2003) constructed a piecewise linear age-depth model based on the dating points.

To resample the times, t^* , for feeding them into the REDFITmc2 algorithm (Sects. 2.7.1 and 2.7.2), we overtake the piecewise model. We impose a constraint to ensure that the $\{t^*(i)\}_{i=1}^n$ increase strictly monotonically.

1. Dating points, $\{z_{\text{date}}(j), t_{\text{date}}(j)\}_{j=1}^{n_{\text{date}}}$.
2. Dating errors, $\{s_{\text{date}}(j)\}_{j=1}^{n_{\text{date}}}$.
3. Timescale, $t(i)=f(z(i))$, $i=1, \dots, n$. The function $f(\cdot)$ is a piecewise linear function through the dating points.
4. Simulated dating points,

$$t_{\text{date}}^*(j)=t_{\text{date}}(j)+s_{\text{date}}(j) \cdot \mathcal{E}_{N(0,1)}(j).$$
5. If $f^*(\cdot)$, the piecewise linear function through the simulated dating points, increases strictly monotonically, then calculate simulated timescale, $t^*(i)=f^*(z(i))$, $i=1, \dots, n$.

3.2.3 Spectrum, with timescale errors

The Q5 spectrum (Fig. 7) exhibits a number of peaks above the upper bounds for the AR(1) hypothesis. Peak I ($T_{\text{period}}=10.9$ a) is significant also when taking the test multiplicity into account. This peak from a Holocene monsoon proxy record may correspond to the cycle found in the

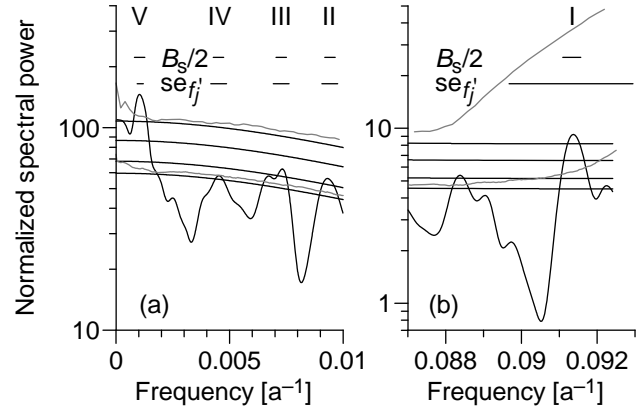


Fig. 7. Spectrum, Q5 $\delta^{18}\text{O}$ record; (a) low-frequency part, (b) high-frequency part. (For a plot over the full frequency range, see Fleitmann et al. (2003, Fig. S5 therein).) Spectrum estimation parameters: OFAC=64, $n_{50}=6$, Welch taper, $B_s \approx 0.001 \text{ a}^{-1}$, $B=10000$ (bias correction). Smooth solid black lines show (from bottom) upper 90%, 95%, 99% and 99.9% chi-squared bounds for an AR(1) red-noise hypothesis ($a'=0.57$); highest bound recognizes the test multiplicity ($M=\text{NINT}[n/(n_{50}+1)]=139$). (Bootstrap bounds (not shown) are nearly identical to chi-squared bounds.) Spectral peaks labelled from I to V, possibly reflecting solar activity variations, are discussed in the text. Their period values are 10.9 a (I), 107 a (II), 137 a (III), 221 a (IV) and 963 a (V). The influence of timescale uncertainties of the Q5 record (Sect. 3.2.2) is studied using algorithm REDFITmc2. The 90% bootstrap bound (the lower of the grey, wiggly lines in a and b) is obtained from mode “a” (Section 2.7.1) with $B=10000$; the 99.9% bound (upper grey lines) with $B=100000$. The frequency uncertainties $se_{f_j'}$ (horizontal bars) due to timescale errors (mode “b”, Sect. 2.7.1), calculated with $B=2000$, are compared with the half of the bandwidth.

record of the number of sunspots, a proxy for solar activity variations, observed since early 17th century (Hoyt and Schatten, 1997). Not as high confidence levels are achieved by the three periods in the centennial band (II, $T_{\text{period}}=107$ a; III, $T_{\text{period}}=137$ a; IV, $T_{\text{period}}=221$ a). The last peak (V, $T_{\text{period}}=963$ a), again strong, may be related to a peak in the spectrum of radiocarbon variations (Stuiver and Braziunas, 1993), which contain information about changes in solar activity.

The timescale of stalagmite Q5 is not exactly known, it has errors stemming from dating uncertainties. How does this influence the detectability and the frequency accuracy of the monsoon peaks?

The upper percentiles of the red-noise alternative obtained with timescale simulations (Fig. 7) are over the whole frequency interval higher than the corresponding percentiles obtained from ignoring dating uncertainties, as expected. This effect seems in case of stalagmite Q5 not excessively large, except for higher frequencies (Fig. 7b). Especially the 99.9% level becomes inflated by the timescale errors, to such a

degree that peak I at $T_{\text{period}} = 10.9$ a is not significant anymore in a multiple test. The only peak in the monsoon spectrum passing the hard test (timescale errors, multiplicity) is that at $T_{\text{period}} = 963$ a.

Feeding the resampled times into the REDFITmc2 algorithm, mode “b” allows to quantify the standard error, $\text{se}_{f_j'}$, of the frequencies of the monsoon peaks owing to timescale errors. Again, the effects are larger on the high-frequency (Fig. 7b) than on the low-frequency side (Fig. 7a). There, the half of the 6-dB bandwidth is in the same order of magnitude as the frequency standard error, $\text{se}_{f_j'}$. On the high-frequency side the error interval for the period of peak I ($T_{\text{period}}=10.9$ a) is from $1/(1/10.9+\text{se}_{f_j'})=10.6$ years to $1/(1/10.9-\text{se}_{f_j'})=11.4$ years.

To summarize, the contribution of spectral analysis to answering the question after the existence of solar peaks in the spectrum of the Holocene monsoon proxy record from stalagmite Q5 is as follows. Peak I corresponds within error bars of frequency to the sunspot cycle, but taking into account timescale errors reduces its multiple test significance considerably. Peaks II ($T_{\text{period}}=107$ a), III ($T_{\text{period}}=137$ a) and IV ($T_{\text{period}}=221$ a), which are partly at periods similar to what is found for the Holocene radiocarbon record (Stuiver and Braziunas, 1993), are not statistically significant (multiple test) even when ignoring timescale errors. Only peak V at $T_{\text{period}}=963$ a, also a solar cycle candidate, is significant.

It would be premature for an analysis of the Sun–monsoon relation to stop at this point. Four lines should be explored. First, the stability of the phasing between both signals should be examined by means of bivariate spectral analysis. Second, the relation can be further investigated, using the same data sets, in the time domain by means of bandpass or harmonic filtering (Ferraz-Mello, 1981). Third, the climate physics of the Sun–monsoon link can be considered. This has been done by Koder (2004), who explained a positive correlation between solar activity and Indian monsoon strength via a weakening of the Brewer–Dobson circulation in the stratosphere. However, this was established on measurement data from 1958–1999, and the feasibility of this or other mechanisms on longer timescales is still elusive. Fourth, other records of Holocene monsoon variations need to be analysed. For example, Neff et al. (2001) analysed a $\delta^{18}\text{O}$ record from a stalagmite from an other cave than where Q5 is from, finding monsoon peaks at $T_{\text{period}}=10.7$ a, 226 a and 1018 a. Combining this evidence with the information from Q5 in a new multiple test should raise the overall statistical significance. A synopsis of evidence pro and contra the Sun–monsoon hypothesis in a multiple statistical test, with timescale errors taken into account, is a major task awaiting to be done.

4 Conclusions

The results from applying the novel REDFITmc2 algorithm for spectrum estimation in the presence of timescale errors support in a quantitative manner the intuition that stronger effects on the spectrum estimate (decreased detectability and increased frequency uncertainty) occur for higher frequencies. Analogously, the results agree with the spectrum distortion (Eq. 9) that is based on assuming timescale errors in the form of independent Gaussian jitter.

The surplus information brought by the algorithm consists of:

1. the corrected upper levels of the spectrum of the red-noise AR(1) alternative, conjectured to be valid for timescale errors of arbitrary form;
2. the uncertainty of the frequency of a spectral peak.

This paper is a first, introductory step. Monte Carlo simulations with pre-defined spectral properties and pre-defined timescale error models, may shed more light on the performance of the REDFITmc2 algorithm for timescale errors of arbitrary form and further support our conjecture. Such simulations may also help to quantify the effects of misspecification of the timescale error model.

Regarding timescale construction, we conclude that not only the fixpoints, dating errors and the functional form of the age–depth model play a role. Also the joint distribution of all time points, that means, serial correlation (jitter) and constraints such as the stratigraphic order, determines spectrum estimation.

Application of REDFITmc2 to ice core and stalagmite records of palaeoclimate allowed a more realistic evaluation of spectral peaks than when ignoring this source of uncertainty.

Acknowledgements. MM acknowledges financial support from the BCF/CACHE program (British Antarctic Survey) and FG 668 of the DFG (Daphne).

Edited by: J. Kurths

Reviewed by: N. Marwan and two anonymous referees

References

- Andersen, K. K., Ditlevsen, P. D., Rasmussen, S. O., Clausen, H. B., Vinther, B. M., Johnsen, S. J., and Steffensen, J. P.: Retrieving a common accumulation record from Greenland ice cores for the past 1800 years, *J. Geophys. Res.*, 111, D15106, doi:10.1029/2005JD006765, 2006.
- Bartlett, M. S.: Periodogram analysis and continuous spectra, *Biometrika*, 37, 1–16, 1950.
- Bendat, J. S. and Piersol, A. G.: *Random Data: Analysis and Measurement Procedures*, 2nd edn. Wiley, New York, 566 pp., 1986.

- Buck, C. E. and Millard, A. R.: Tools for Constructing Chronologies: Crossing Disciplinary Boundaries. Springer, London, 257 pp., 2004.
- Chan, K. S. and Tong, H.: A note on embedding a discrete parameter ARMA model in a continuous parameter ARMA model, *J. Time Ser. Anal.*, 8, 277–281, 1987.
- Divine, D. V., Polzehl, J., and Godtlielsen, F.: A propagation-separation approach to estimate the autocorrelation in a time-series, *Nonlin. Processes Geophys.*, 15, 591–599, 2008, <http://www.nonlin-processes-geophys.net/15/591/2008/>.
- Efron, B. and Tibshirani, R. J.: *An Introduction to the Bootstrap*. Chapman and Hall, London, 436 pp., 1993.
- EPICA Community Members: One-to-one coupling of glacial climate variability in Greenland and Antarctica, *Nature*, 444, 195–198, 2006.
- Ferraz-Mello, S.: Estimation of periods from unequally spaced observations, *Astronom. J.*, 86, 619–624, 1981.
- Fisher, D. A., Reeh, N., and Clausen, H. B.: Stratigraphic noise in time series derived from ice cores, *Ann. Glaciol.*, 7, 76–83, 1985.
- Fleitmann, D., Burns, S. J., Mudelsee, M., Neff, U., Kramers, J., Mangini, A., and Matter, A.: Holocene forcing of the Indian monsoon recorded in a stalagmite from Southern Oman, *Science*, 300, 1737–1739, 2003.
- Fleitmann, D., Burns, S. J., Mangini, A., Mudelsee, M., Kramers, J., Villa, I., Neff, U., Al-Subbary, A. A., Buettner, A., Hippler, D., and Matter, A.: Holocene ITCZ and Indian monsoon dynamics recorded in stalagmites from Oman and Yemen (Socotra), *Quat. Sci. Rev.*, 26, 170–188, 2007.
- Fleitmann, D., Mudelsee, M., Burns, S. J., Bradley, R. S., Kramers, J., and Matter, A.: Evidence for a widespread climatic anomaly at around 9.2 ka before present, *Paleoceanog.*, 23, PA1102, doi:10.1029/2007PA001519, 2008.
- Harris, F. J.: On the use of windows for harmonic analysis with the discrete Fourier transform, *Proc. IEEE*, 66, 51–83, 1978.
- Heegaard, E., Birks, H. J. B., and Telford, R. J.: Relationships between calibrated ages and depth in stratigraphical sequences: An estimation procedure by mixed-effect regression, *The Holocene*, 15, 612–618, 2005.
- Horne, J. H. and Baliunas, S. L.: A prescription for period analysis of unevenly sampled time series, *Astrophys. J.*, 302, 757–763, 1986.
- Hoyt, D. V. and Schatten, K. H.: *The Role of the Sun in Climate Change*. Oxford University Press, New York, 279 pp., 1997.
- Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, A., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and millennial Antarctic climate variability over the past 800 000 years, *Science*, 317, 793–796, 2007.
- Kendall, M. G.: Note on bias in the estimation of autocorrelation, *Biometrika*, 41, 403–404, 1954.
- Kodera, K.: Solar influence on the Indian Ocean monsoon through dynamical processes, *Geophys. Res. Lett.*, 31, L24209, doi:10.1029/2004GL020928, 2004.
- Lisiecki, L. E. and Raymo, M. E.: A Plio-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records, *Paleoceanog.*, 20, PA1003, doi:10.1029/2004PA001071, 2005.
- Lomb, N. R.: Least-squares frequency analysis of unequally spaced data, *Astro. Sp. Sc.*, 39, 447–462, 1976.
- Moore, M. I. and Thomson, P. J.: Impact of jittered sampling on conventional spectral estimates, *J. Geophys. Res.*, 96, 18519–18526, 1991.
- Mudelsee, M.: Ramp function regression: A tool for quantifying climate transitions, *Comput. Geosci.*, 26, 293–307, 2000.
- Mudelsee, M.: Note on the bias in the estimation of the serial correlation coefficient of AR(1) processes, *Stat. Pap.*, 42, 517–527, 2001.
- Mudelsee, M.: TAUEST: A computer program for estimating persistence in unevenly spaced weather/climate time series, *Comput. Geosci.*, 28, 69–72, 2002.
- Neff, U., Burns, S. J., Mangini, A., Mudelsee, M., Fleitmann, D., and Matter, A.: Strong coherence between solar variability and the monsoon in Oman between 9 and 6 kyr ago, *Nature*, 411, 290–293, 2001.
- Nuttall, A. H.: Some windows with very good sidelobe behavior, *IEEE T. Acoust. Speech*, 29, 84–91, 1981.
- Parrenin, F., Barnola, J.-M., Beer, J., Blunier, T., Castellano, E., Chappellaz, J., Dreyfus, G., Fischer, H., Fujita, S., Jouzel, J., Kawamura, K., Lemieux-Dudon, B., Loulergue, L., Masson-Delmotte, V., Narcisi, B., Petit, J.-R., Raisbeck, G., Raynaud, D., Ruth, U., Schwander, J., Severi, M., Spahni, R., Steffensen, J. P., Svensson, A., Udisti, R., Waelbroeck, C., and Wolff, E.: The EDC3 chronology for the EPICA Dome C ice core, *Clim. Past*, 3, 485–497, 2007.
- Priestley, M. B.: *Spectral Analysis and Time Series*, Academic Press, London, 890 pp., 1981.
- Raymo, M. E. and Huybers, P.: Unlocking the mysteries of the ice ages, *Nature*, 451, 284–285, 2008.
- Robinson, P. M.: Estimation of a time series model from unequally spaced data, *Stoch. Process. Appl.*, 6, 9–24, 1977.
- Röthlisberger, R., Mudelsee, M., Bigler, M., de Angelis, M., Fischer, H., Hansson, M., Lambert, F., Masson-Delmotte, V., Sime, L., Udisti, R., and Wolff, E. W.: The Southern Hemisphere at glacial terminations: insights from the Dome C ice core, *Clim. Past*, 4, 345–356, 2008, <http://www.clim-past.net/4/345/2008/>.
- Scargle, J. D.: Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data, *Astrophys. J.*, 263, 835–853, 1982.
- Scargle, J. D.: Studies in astronomical time series analysis. III. Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data, *Astrophys. J.*, 343, 874–887, 1989.
- Schulz, M. and Stattegger, K.: SPECTRUM: Spectral analysis of unevenly spaced paleoclimatic time series, *Comput. Geosci.*, 23, 929–945, 1997.
- Schulz, M. and Mudelsee, M.: REDFIT: Estimating red-noise spectra directly from unevenly spaced paleoclimatic time series, *Comput. Geosci.*, 28, 421–426, 2002.
- Schuster, A.: On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena, *Terrestr. Magn.*, 3, 13–41, 1898.
- Spötl, C., Scholz, D., and Mangini, A.: A terrestrial U/Th-dated stable isotope record of the penultimate interglacial, *Earth Planet. Sci. Lett.*, 276, 283–292, 2008.

- Stuiver, M. and Braziunas, T. F.: Sun, ocean, climate and atmospheric $^{14}\text{CO}_2$: An evaluation of causal and spectral relationships, *Holocene*, 3, 289–305, 1993.
- Thomson, D. J.: Spectrum estimation and harmonic analysis, *Proc. IEEE*, 70, 1055–1096, 1982.
- Thomson, P. J. and Robinson, P. M.: Estimation of second-order properties from jittered time series, *Ann. Inst. Statist. Math.*, 48, 29–48, 1996.
- VanDongen, H. P. A., Olofsen, E., VanHartevelt, J. H., and Kruij, E. W.: Periodogram analysis of unequally spaced data: The Lomb method. Leiden University, Leiden, 66 pp., 1997.
- Welch, P. D.: The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE T. Audio Electro.*, 15, 70–73, 1967.
- Wolff, E. W., Fischer, H., Fundel, F., Ruth, U., Twarloh, B., Littot, G. C., Mulvaney, R., Röthlisberger, R., de Angelis, M., Boutron, C. F., Hansson, M., Jonsell, U., Hutterli, M. A., Lambert, F., Kaufmann, P., Stauffer, B., Stocker, T. F., Steffensen, J. P., Bigler, M., Siggaard-Andersen, M. L., Udisti, R., Becagli, S., Castellano, E., Severi, M., Wagenbach, D., Barbante, C., Gabrielli, P., and Gaspari, V.: Southern Ocean sea-ice extent, productivity and iron flux over the past eight glacial cycles, *Nature*, 440, 491–496, 2006.
- Wunsch, C.: On sharp spectral lines in the climate record and the millennial peak, *Paleoceanog.*, 15, 417–424, 2000.